SRB STATES AND NONEQUILIBRIUM STATISTICAL MECHANICS CLOSE TO EQUILIBRIUM.

by Giovanni Gallavotti* and David Ruelle**

Abstract. *Nonequilibrium statistical mechanics close to equilibrium is studied using SRB states and a formula [10] for their derivatives with respect to parameters. We write general expressions for the thermodynamic fluxes (or currents) and the transport coefficients, generalizing the results of [4], [5]. In this framework we give a general proof of the Onsager reciprocity relations.*

There is currently a strong revival of nonequilibrium statistical mechanics, based on taking seriously the nonlinear (chaotic) microscopic dynamics. One natural idea in this direction is to use nonequilibrium steady states, which are defined to be the SRB states for the dynamics (see below).

The idea of using SRB states eventually led to useful results only recently, when it was combined in [6] with reversibility of the dynamics to obtain a nontrivial fluctuation formula for the entropy production. In [5] the same ideas were applied to prove the Onsager reciprocity relations, and the fluctuation-dissipation formula for a rather special class of models. The analysis dealt with examples rather than the general situation, and relied on an unproven conjecture on Anosov systems. In this note we generalize [5] and, using [10], give results that can be proved rigorously for *Axiom A* diffeomorphisms (this is a strong chaoticity assumption, see [12], here we skip technical details).

To be discussed is the classical microscopic description of the time evolution of a physical system. (No *large system* assumption will be made). We let the accessible phase space be a compact manifold M. The time evolution is given by iterates of a diffeomorphism f of M (discrete time case) or by integrating a vector field F on M (continuous time case). We choose a Riemann metric on M, and let dx be the corresponding volume element.

In nonequilibrium statistical mechanics, the time evolution typically does not preserve any measure which has a density with respect to dx. Let $m(dx) = \underline{m}(x)dx$ be a probability measure (with density $\underline{m}$), and $f^{*k}m$ the direct image of m by f^{*k}; any weak limit for $n \to \infty$ of

* Permanent address: Fisica, Universitá di Roma “La Sapienza”, 00185 Roma, Italia
** Permanent address: IHES, 91440 Bures sur Yvette, France.

$$\frac{1}{n}\sum_{k=0}^{n-1} f^{*k} m$$

is an f-invariant probability measure ρ_f on M. If furthermore ρ_f is ergodic we may say that it is a *natural nonequilibrium state* (the SRB states are special examples of this, [3]).

An infinitesimal change δf of f corresponds to an (infinitesimal) vector field $X = \delta f \circ f^{-1}$ on M and an easy formal calculation gives

$$\delta\rho_f(\Phi) = \sum_{n=0}^{\infty} \rho_f \langle \mathrm{grad}(\Phi \circ f^n), X\rangle \tag{1}$$

where Φ is a smooth test function and $\langle \cdot\, , \cdot \rangle$ is the scalar product of a tangent vector and a cotangent vector to M. Equation (1) expresses the change in the expectation value of the observable Φ when the system is subjected to a force X. Replacing X by a time dependent force X_t we have for the change of expectation value of the observable Φ at time s the analogous formula

$$\delta_s\rho_f(\Phi) = \sum_{t \le s} \rho_f \langle \mathrm{grad}(\Phi \circ f^{-t}), X_t\rangle \tag{2}$$

Note that the condition $t \le s$ corresponds to the "causality principle" and (2) can be used to derive Kramers-Kronig dispersion relations. For Axiom A diffeomorphisms, (1) and (2) can be proved rigorously*: the left-hand side is a derivative, and the right-hand side a convergent series.

In the case of continuous time systems described by a differential equation $\dot{x} = F(x)$ and by the corresponding flow (f^t), an infinitesimal variation $\delta F = X$ generates a variation in the expectation value $\rho_F(\Phi)$ (for the *natural nonequilibrium state*) given by:

$$\delta\rho_F(\Phi) = \int_0^{+\infty} dt \int \rho_F(dx) \langle\, \mathrm{grad}_x(\Phi \circ f^t), X(x)\rangle \tag{1'}$$

(A rigorous proof for Axiom A flows has not been given yet).

The entropy production associated with the diffeomorphism f is defined by**

$$e_f = \rho_f(\sigma_f) \qquad , \qquad \sigma_f = -\log J_f \tag{3}$$

where J_f is the absolute value of the Jacobian of f with respect to the Riemann volume element dx. In the continuous time systems, we let

$$e_F = \rho_F(\sigma_F) \qquad , \qquad \sigma_F = -\mathrm{div} F \tag{3'}$$

(see [5]).

* See [10]; the proof given in [10] for the time dependent case assumes that the pertubation has finite support in time.

** For a discussion of entropy production see [9].

From now on we shall fix f such that $\rho_f(dx) = dx$, *i.e.*, ρ_f is the Riemann volume element (if ρ_f has smooth density this can be achieved by a change of metric). Note that in particular this implies that (1) can be rewritten as

$$\delta\rho_f(\Phi) = \sum_{n=0}^{\infty} \rho_f(\Phi \circ f^n . (-\mathrm{div} X)) \tag{4}$$

In this *nondissipative* situation we have $\log J_f = 0$, hence $e_f = 0$. If we write as before $X = \delta f \circ f^{-1}$ we obtain to second order in X, using (3) and (1),

$$\begin{aligned} e_{f+\delta f} &= (\rho_f + \delta\rho_f)(\sigma_{f+\delta f}) \\ &= \frac{1}{2}\rho_f((\mathrm{div}X)^2) - \delta\rho_f(\mathrm{div}X \circ f) \\ &= \frac{1}{2}\rho_f((\mathrm{div}X)^2) - \sum_{n=0}^{\infty} \rho_f \langle \mathrm{grad}((\mathrm{div}X) \circ f^n), X\rangle \\ &= \frac{1}{2}\rho_f((\mathrm{div}X)^2) + \sum_{n=1}^{\infty} \rho_f((\mathrm{div}X) \circ f^n . \, \mathrm{div}X) \qquad (5) \\ &= \frac{1}{2}\sum_{n=-\infty}^{\infty} \rho_f((\mathrm{div}X) \circ f^n . \, \mathrm{div}X) \end{aligned}$$

The same analysis leads, in the continuous time case (with $\delta F = X$, and $\rho_F(dx) = dx$) to:

$$\begin{aligned} e_{F+\delta F} &= \frac{1}{2}\int_{-\infty}^{+\infty} dt\, \rho_f(\sigma_f \circ f^t \cdot \sigma_f) \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} dt \int dx \ \mathrm{div}\, X(f^t x) \ \cdot \ \mathrm{div}\, X(x) \qquad (5') \end{aligned}$$

see [4], [5].

We shall now relate the above expressions for the entropy production to the definition of the *thermodynamic forces* $\mathcal{X}_\alpha$, and the *conjugated thermodynamic fluxes* $\mathcal{J}_\alpha$ as they appear for instance in [8]. We begin by an informal discussion, and assume, as usual in applications, that f or F depends on parameters E_α, so that we may write (to first order)

$$X = \sum_\alpha V_\alpha \delta E_\alpha$$

We identify the thermodynamic forces $\mathcal{X}_\alpha$ with the parameters E_α. Considering first the continuous time case, we follow [4], [5], and we define the thermodynamic flux conjugated to E_α as

$$\mathcal{J}_\alpha = \rho_{F+\delta F}(\frac{\partial}{\partial E_\alpha}\sigma_{F+\delta F})$$

Since $\partial\sigma/\partial E_\alpha$ is a divergence, $\rho(\partial\sigma/\partial E_\alpha) = 0$ and we have

$$\mathcal{J}_\alpha = \delta\rho_F(-\mathrm{div}V_\alpha) \qquad + \text{h.o.}$$

$$= \int_0^{+\infty} dt \int \rho_F(dx)\langle \mathrm{grad}_x((-\mathrm{div}V_\alpha)\circ f^t), X(x)\rangle \qquad + \text{h.o.}$$

From now on we neglect higher order terms and (using integration by parts, since $\rho_F(dx) = dx$) we write

$$\mathcal{J}_\alpha = \int_0^{+\infty} dt \int \rho_F(dx)\,(\mathrm{div}_x X)(\mathrm{div}_{f^t x} V_\alpha)$$

In the discrete time case we define the thermodynamic flux only to leading order in the E_α by

$$\mathcal{J}_\alpha = \frac{1}{2}\rho_f((\mathrm{div}V_\alpha)(\mathrm{div}X)) + \sum_{n=1}^{\infty}\rho_f((\mathrm{div}V_\alpha)\circ f^n.\,(\mathrm{div}X))$$

With these definitions $\mathcal{J}_\alpha$ depends only on the application of X in the past (causality, cf (2)) and the entropy production (to second order) is

$$e_{f+\delta f} \qquad \text{or} \qquad e_{F+\delta F} = \sum_\alpha \mathcal{X}_\alpha \mathcal{J}_\alpha$$

These conditions uniquely determine the $\mathcal{J}_\alpha$. Notice that the formulae for $\mathcal{J}_\alpha$ involve only the divergences of X and V_α.

To continue the discussion, we assume that there is a (sufficiently large) Banach space $\mathcal{B}$ of functions $\Phi : M \to \mathbf{R}$ such that

$$\rho_f(\Phi) = 0 \qquad\qquad \text{if} \qquad\qquad \Phi \in \mathcal{B}$$

and for some constant C

$$\sum_{k\in\mathbf{Z}} |\rho_f(\Psi\circ f^k.\,\Phi)| \ \le C\|\Phi\|_{\mathcal{B}}\|\Psi\|_{\mathcal{B}} \qquad\qquad \text{if} \qquad\qquad \Phi,\Psi\in\mathcal{B} \tag{6}$$

[This is the discrete time case, the continuous time case is similar. If f is an Anosov diffeomorphism* we can take for $\mathcal{B}$ a space of Hölder continuous functions on M. Similarly for Anosov flows].

From now on, we assume that $\mathrm{div}X$ is in the Banach space $\mathcal{B}$ just introduced, and we may define $\mathcal{X} \in \mathcal{B}$ and $\mathcal{J} \in \mathcal{B}^*$ (the dual of $\mathcal{B}$) as follows:

$$\mathcal{X} = -\mathrm{div}X \in \mathcal{B} \tag{7}$$

$$(\mathcal{J},\Phi) = \frac{1}{2}\rho_f(-\mathrm{div}X.\,\Phi) + \sum_{n=1}^{\infty}\rho_f((-\mathrm{div}X)\circ f^n.\,\Phi) \tag{8}$$

* An Axiom A diffeomorphism f preserving dx is an Anosov diffeomorphism.

for discrete time, and

$$(\mathcal{J},\Phi)=\frac{1}{2}\int_0^\infty \rho_F((-\mathrm{div}X)\circ f^t.\,\Phi) \tag{8'}$$

for continuous time, where $\Phi\in\mathcal{B}$, and $(\cdot\,,\cdot)$ is the pairing $\mathcal{B}^*\times\mathcal{B}\to\mathbf{C}$. Note that if $\mathrm{div}V_\alpha\in\mathcal{B}$ we have

$$\mathcal{J}_\alpha=(\mathcal{J},-\mathrm{div}V_\alpha)$$

and that the entropy production is

$$e_{f+\delta f}\qquad\text{or}\qquad e_{F+\delta F}=(\mathcal{J},\mathcal{X})$$

With the above notation and assumptions we may write $\mathcal{J}=L\mathcal{X}$, where L is, in view of (6), a continuous linear map $\mathcal{B}\to\mathcal{B}^*$. If we define a unitary operator U on $L^2(\rho_f)$ by $U\Phi=\Phi\circ f$ and write $U=\int_{-\pi}^{\pi}e^{i\alpha}\,d\mathbf{P}(\alpha)$ we have

$$\sum_{k\in\mathbf{Z}}e^{-ik\alpha}\rho_f(\Phi\circ f^k.\,\Phi)=2\pi\frac{d}{d\alpha}(\Phi,d\mathbf{P}(\alpha)\Phi)_{L^2}$$

so that the quadratic form associated with L satisfies

$$(L\Phi,\Phi)=\frac{1}{2}\sum_{k\in\mathbf{Z}}\rho_f(\Phi\circ f^k.\,\Phi)=\pi\frac{d}{d\alpha}(\Phi,d\mathbf{P}(\alpha)\Phi)_{L^2}\Big|_{\alpha=0}\geq 0$$

In particular, this quadratic form is ≥ 0.

The formulae obtained up to now hold under the only assumption of closeness to equilibrium. If we make the further assumption of (microscopic) reversibility, we shall obtain a symmetry property of L called Onsager reciprocity. For simplicity we discuss only the discrete time case.

We say that the dynamics is *reversible* if there exists a diffeomorphism $i:M\to M$ such that $i^2=$ identity, $i\circ f=f^{-1}\circ i$. We have then also $i^*\rho_f=\rho_f$. [Note that ρ_f, *i.e.*, the Riemann volume, is mixing by (6), hence f-ergodic. Since $i^*\rho_f$ is absolutely continuous with respect to ρ_f and satisfies $f^*(i^*\rho_f)=i^*(f^{-1})^*\rho_f=(i^*\rho_f)$ we have $i^*\rho_f=\rho_f$ by ergodicity]. Assuming reversibility we may define $\epsilon\Phi=\Phi\circ i$ for $\Phi\in\mathcal{B}$, and we find

$$(L\Psi,\Phi)=\frac{1}{2}\rho_f(\Psi.\,\Phi)+\sum_{n=1}^{\infty}\rho_f(\Psi\circ f^n.\,\Phi)=\frac{1}{2}\rho_f(\epsilon\Psi.\,\epsilon\Phi)+\sum_{n=1}^{\infty}\rho_f(\epsilon\Psi\circ f^{-n}.\,\epsilon\Phi)$$

$$=\frac{1}{2}\rho_f(\epsilon\Phi.\,\epsilon\Psi)+\sum_{n=1}^{\infty}\rho_f(\epsilon\Phi\circ f^{-n}.\,\epsilon\Psi)=(L(\epsilon\Phi),\epsilon\Psi)$$

The relation

$$(L\Psi,\Phi)=(L(\epsilon\Phi),\epsilon\Psi)$$

is a form of the *Onsager reciprocity relation* as we shall see in a moment. Note that reversibility was assumed only for f (*i.e.*, at equilibrium), the perturbation δf is arbitrary.

To obtain a more familiar form of the entropy production formula (see [8]), we assume that $\mathcal{B}$ has a basis (Φ_α) with a corresponding system (ϕ_α) of elements of $\mathcal{B}^*$ such that $(\phi_\alpha, \Phi_\beta) = \delta_{\alpha\beta}$ (see [11]). Write

$$\mathcal{X}_\alpha = (\phi_\alpha, \mathcal{X}) = (\phi_\alpha, -\text{div}X)$$

$$\mathcal{J}_\alpha = (\mathcal{J}, \Phi_\alpha) = \frac{1}{2}\rho_f(-\text{div}X, \Phi_\alpha) + \sum_{n=0}^{\infty} \rho_f((-\text{div}X)\circ f^n . \Phi_\alpha)$$

In particular

$$-\text{div}X = \sum_\alpha (\phi_\alpha, -\text{div}X).\, \Phi_\alpha = \sum_\alpha \mathcal{X}_\alpha .\, \Phi_\alpha$$

and

$$e_{f+\delta f} = (\mathcal{J}, \mathcal{X}) = \sum_\alpha \mathcal{J}_\alpha .\, \mathcal{X}_\alpha$$

$$\mathcal{J}_\alpha = \frac{1}{2}\sum_\beta \mathcal{X}_\beta \rho_f(\Phi_\beta \Phi_\alpha) + \sum_{n=1}^{\infty}\sum_\beta \mathcal{X}_\beta \rho_f(\Phi_\beta \circ f^n .\, \Phi_\alpha)$$

To avoid convergence problems suppose that finitely many $\mathcal{X}_\beta$ only are nonzero. Then

$$\mathcal{J}_\alpha = \sum_\beta L_{\alpha\beta}\mathcal{X}_\beta$$

where

$$L_{\alpha\beta} = \frac{1}{2}\rho_f(\Phi_\beta \Phi_\alpha) + \sum_{n=1}^{\infty} \rho_f(\Phi_\beta \circ f^n .\, \Phi_\alpha)$$

Suppose again reversibility of the dynamics, and let the basis (Φ_α) of $\mathcal{B}$ be such that $\epsilon\Phi_\alpha = \Phi_\alpha \circ f = \epsilon_\alpha \Phi_\alpha$ with $\epsilon_\alpha = \pm 1$. Then

$$L_{\alpha\beta} = \epsilon_\alpha \epsilon_\beta L_{\beta\alpha}$$

which is the usual form of the Onsager reciprocity relation.

We conclude by sketching an example, see [1], of the formalism described above. Let Σ be a surface of constant negative curvature, and genus g, with the $g-1$ automorphic forms $\phi_\alpha(z)dz$. We normalize these forms so that they are orthonormal in the space $L_2(T\Sigma)$ in the natural scalar product, [7].

We can then consider the hamiltonian equations of the motion of a particle on Σ subject to the external force generated by the "electric" field $\mathcal{E}$ such that $\mathcal{E}_x dx + \mathcal{E}_y dy = \text{Re} \sum E_\alpha \varphi_\alpha(z) dz$.

We also impose, via Gauss' principle, [5], that there is a *thermostat* force that keeps the kinetic energy constant (and equal to 1/2) in spite of the field's action. Thus the "only" efffect that the fields will have on the flow is that currents flowing "around" the g "holes" of the surface will be established. But the flow being a mixing Anosov flow on a compact surface (because of the gaussian constraint) it will result that a stationary state will be reached and the latter will be the SRB distribution, [2].

The equations of motion can be written explicitly and one finds in particular that the entropy creation rate at the point $\vec{q}, \vec{p}$ of the phase space is

$$\sigma(\vec{q}, \vec{p}) = \frac{\vec{\mathcal{E}} \cdot \vec{p}}{p^2}$$

The transport coefficient can also be explicitly computed, and one finds $L_{\alpha\beta} = \frac{1}{2}\delta_{\alpha\beta}$, see [1].

References.

[1] F.Bonetto, G.Gentile, V.Mastropietro. "Electric fields on a surface of constant negative curvature." Preprint, 1996.

[2] R.Bowen, D.Ruelle. "Ergodic theory of Axiom A flows." Inventiones Math. **29**,181-202(1975).

[3] J.-P.Eckmann and D.Ruelle. "Ergodic theory of chaos and strange attractors" Rev. Mod. Phys. **57**,617-656(1985).

[4] G.Gallavotti: "Extension of Onsager's reciprocity to large fields and the chaotic hypothesis." Phys. Rev. Letters. **77**,4334-4337(1996).

[5] G.Gallavotti. "Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem." J. Statist. Phys. **84**, 899–926(1996)

[6] G.Gallavotti and E.G.D.Cohen. "Dynamical ensembles in nonequilibrium statistical mechanics." Phys. Rev. Letters **74**,2694-2697(1995). "Dynamical ensembles in stationary states." J. Statist. Phys. **80**,931-970(1995).

[7] I.M.Gel'fand, M.I.Graev. I.I.Pyateckii-Shapiro. *Representation theory and automorphic functions.* Saunders, Philadelphia, 1969.

[8] S.R.de Groot and P.Mazur. *Non-equilibrium statistical thermodynamics.* Dover, New York, 1984.

[9] D.Ruelle. "Positivity of entropy production in nonequilibrium statistical mechanics." J. Statist. Physics, **85**,1-25(1996).

[10] D.Ruelle. "Differentiation of SRB states." To be published; the preprint is archived in `mp_arc@math.utexas.edu`, **96**, #499.

[11] I.Singer. *Bases in Banach spaces I.* Springer, Berlin, 1970.

[12] S.Smale. "Differentiable dynamical systems." Bull. Amer. Math. Soc. **73**,747-817(1967).